Scatter Correction for Positron Emission Mammography

Jinyi Qi and Ronald H Huesman

Abstract—**We have previously presented a regularized list mode maximum likelihood reconstruction algorithm for the positron emission tomograph that is being developed at our laboratory. Here we will present a scatter correction method for this algorithm. The mean scatter sinogram is estimated using a Monte Carlo simulation program. It is then incorporated into the forward model of the reconstruction algorithm. With the assumption that the background activity is nearly uniform, the Monte Carlo scatter simulation need only run once for each PEM configuration. This saves computational time and makes the Monte Carlo scatter correction viable in clinical situations. The propagation of the noise from the estimated scatter sinogram into the reconstruction is theoretically analyzed. The results provide an easy way to calculate the required number of events in the Monte Carlo scatter simulation for a given noise level in the image. The analysis is also applicable to other scatter estimation methods provided that the covariance of the estimated scatter sinogram is available.**

I. Introduction

A rectangular positron emission tomograph (Fig. 1), dedicated to imaging the human breast, is under development at our Laboratory [1]. The tomograph consists of four banks of detector modules (two banks of 3×3 modules left and right and two banks of 3×4 modules top and bottom). Each module consists of an 8×8 array of 3mm×3mm×30mm lutetium oxyorthosilicate (LSO) crystals. The maximum field of view (FOV) of the system is 96×72×72 mm^3. For each crystal, the system digitizes the depth of interaction (DOI) of the photon with three bits. Each detector is placed in coincidence with all detectors in the other three banks, giving rise to 172 million possible lines of response (LORs). The system operates exclusively in fully 3D mode.

The data from the new tomograph is stored in list mode format because the total number of detections will generally be far less than the total number of LORs. We have developed a list mode likelihood reconstruction algorithm for the tomograph [2]. The DOI information was explicitly modeled in the forward projection for each LOR.

Here we present a scatter correction method for the list mode likelihood reconstruction algorithm. The scatter sinogram is estimated using a Monte Carlo simulation program. The results are then incorporated in the forward model of the reconstruction algorithm. For breast imaging with F-18-labeled deoxyglucose (FDG), we can assume that the FOV is filled with uniform activity and that features such as cancerous lesions account for a small fraction of the radioactivity, so we only need to run the Monte Carlo scatter simulation once for each system configuration. For each data set, we can scale the scatter sinogram by the ratio of the total events between the Monte Carlo simulation and the data set, assuming the scatter profiles are the same. This saves a large amount of computation time. We also theoretically analyzed the noise propagation from the Monte Carlo scatter sinogram into the reconstructed image. Such analysis is also useful in determining the total number of events that is required for the Monte Carlo scatter simulation. Using Monte Carlo simulation to estimate scatter sinogram is not a new idea. It is, however, the application of the method to the PEM reconstruction and the theoretical analysis of the noise propagation that make this paper novel.

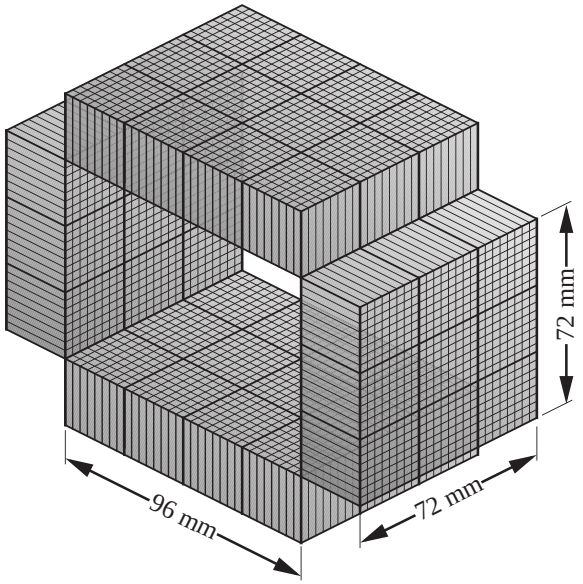

Fig. 1. PEM Geometry.

II. Theory

A. List Mode Likelihood Reconstruction with Scatter Correction

Histogrammed PET data are generally modeled as a collection of independent Poisson random variables. By treating the detections in each LOR separately, we can derive the appropriate log-likelihood function for list mode data [2]:

$$L(\boldsymbol{x}) = \sum_{k=1}^{K} \log \left[\sum_{j=1}^{N} p(i_k, j) x_j + s_{i_k} \right] - \sum_{j=1}^{N} \varepsilon_j x_j, \qquad (1)$$

This work was supported by the U.S. Department of Health and Human Services under grant P01 HL25840, by the National Cancer Institute under grant R01 CA 59794, and by the Director, Office of Science, Office of Biological and Environmental Research, Medical Sciences Division of the US Department of Energy under contract DE-AC03-76SF00098.

J. Qi and R.H. Huesman are with the Center for Functional Imaging, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA (telephone: 510-486-4695, e-mail: {jqi,rhhuesman}@lbl.gov).

where x_j is mean activity inside the j^{th} voxel of the unknown image, $p(i,j)$ is the probability of detecting an event from the j^{th} voxel in the i^{th} LOR, s_i is the mean scatter in the i^{th} LOR, i_k is the index of the LOR of the k^{th} detection, $\varepsilon_j \equiv \sum_i p(i,j)$, K is the total number of detections, and N is the total number of image voxels.

The maximum likelihood (ML) estimate can be found by maximizing (1). A popular ML algorithm for PET reconstruction is the expectation maximization (EM) algorithm [3], [4], [5]. However, the ML solution is unstable (i.e. noisy) because of the ill-conditioness of tomography systems. Hence some form of regularization (or prior function) is needed to reconstruct a reasonable image. The prior function used in [2] is a Gaussian prior whose logarithm is of the form

$$\beta U(\boldsymbol{x}) = \frac{\beta}{2}(\boldsymbol{x}-\boldsymbol{m})'\boldsymbol{R}(\boldsymbol{x}-\boldsymbol{m}), \tag{2}$$

where β is the smoothing parameter, $\boldsymbol{m}$ is the estimated mean of the unknown image, $\boldsymbol{R}$ is a positive definite (or semidefinite) matrix.

Combining the likelihood function (1) and the image prior (2), the reconstruction is found as:

$$\hat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x}\geq 0} \left[L(\boldsymbol{x}) - \beta U(\boldsymbol{x})\right]. \tag{3}$$

For further simplification, $\boldsymbol{R}$ is chosen to be a diagonal matrix, so the EM algorithm can be used to solve (3). The EM update equation is [2]

$$\hat{x}_j^{n+1} = \frac{1}{2}\left(m_j - \frac{\varepsilon_j}{\beta r_{jj}}\right) + \sqrt{\frac{1}{4}\left(m_j - \frac{\varepsilon_j}{\beta r_{jj}}\right)^2 + \frac{\hat{x}_j^n}{\beta r_{jj}} \sum_{k=1}^{K} \frac{p(i_k,j)}{\sum_{l=1}^{N} p(i_k,l)\hat{x}_l^n + s_{i_k}}},$$

where r_{jj} is the (j,j)th element of $\boldsymbol{R}$.

B. Estimate Mean Scatter Sinogram using Monte Carlo Simulation

The scatter correction method described in the previous section requires the mean scatter sinogram being known before the reconstruction starts. For conventional PET systems, a scatter sinogram can be estimated by single scatter simulation, deconvolution of emission sinogram, dual energy windows, or Monte Carlo methods. Most analytical scatter estimation methods require fitting an computed scatter sinogram to the tails of the emission sinogram that consist of pure scatter events. This is not practical for PEM geometry as the whole FOV is filled with activity. Therefore, we adopt the Monte Carlo method here.

The Monte Carlo scatter simulation requires both emission and attenuation maps. Generally they are obtained from an initial reconstruction without scatter correction. One advantage of breast imaging with FDG is that the background is quite uniform. If we assume that whole FOV is filled with uniform activity and that features such as cancerous lesions account for a small fraction of the radioactivity, then we only need to run the Monte Carlo simulation once for each system configuration. This will saves a large amount of computational time.

The Monte Carlo simulation program estimates the scatter sinogram by tracing all 511 keV photon pairs randomly generated inside the FOV. For each photon, it first computes the interaction point based on the attenuation length. Then, it determines whether it is a photo-electric or Compton interaction. If it is photo-electric, it dumps all of its current energy; if it is Compton, it computes the energy deposited and the new direction of the photon using the Klein-Nishima formula, and continues to trace the photon until the photon has dumped all its energy or traveled outside of the system. A photon is detected when the energy deposited at one detector is greater than a preselected threshold. A coincidence event is recorded if both photons are detected. The Monte Carlo simulation program histograms separately the scattered and unscattered (true) events.

For each individual data set, the scatter sinogram is then estimated by

$$\hat{s}_i = \frac{\text{total events in the data}}{\text{total events in Monte Carlo Simulation}} s_i^{MC}, \tag{4}$$

where s_i^{MC} is the number of scatter events in the i^{th} LOR in the Monte Carlo simulation. This assumes the scatter fraction and profile in each data set are the same, as most scatters are generated from the uniform background.

C. Noise Propagation

Noise is inevitable in the Monte Carlo scatter sinogram. Because the Poisson nature of the counting process, the longer it runs, the less the noise is. This presents a trade off between time and accuracy. In this section, we analyze how the noise in the scatter sinogram propagates into the reconstruction.

We denote the MAP reconstruction in (3) as $\hat{\boldsymbol{x}}(\boldsymbol{y}, \hat{\boldsymbol{s}})$ to indicate that $\hat{\boldsymbol{x}}$ is dependent on estimated scatter sinogram $\hat{\boldsymbol{s}}$. Since $\boldsymbol{y}$ and $\hat{\boldsymbol{s}}$ are independent, we first focus on noise from $\hat{\boldsymbol{s}}$ and assume $\boldsymbol{y}$ noisefree (i.e. $\boldsymbol{y} = \bar{\boldsymbol{y}} = \boldsymbol{P}\boldsymbol{x} + \boldsymbol{s}$). We can approximate $\hat{\boldsymbol{x}}(\boldsymbol{y}, \hat{\boldsymbol{s}})$ using a first order Taylor series expansion at the point $\hat{\boldsymbol{s}} = \boldsymbol{s}$:

$$\hat{\boldsymbol{x}}(\boldsymbol{y}, \hat{\boldsymbol{s}}) \approx \hat{\boldsymbol{x}}(\boldsymbol{y}, \boldsymbol{s}) + \nabla_s \hat{\boldsymbol{x}}(\boldsymbol{y}, \boldsymbol{s})(\hat{\boldsymbol{s}} - \boldsymbol{s}). \tag{5}$$

This approximation is similar to that presented in [6]. From (5), we have the following expression for the covariance of noise in the reconstruction caused by the noise in the estimated scatter sinogram

$$\Sigma(\hat{\boldsymbol{x}}) \approx \nabla_s \hat{\boldsymbol{x}}(\boldsymbol{y}, \boldsymbol{s}) \Sigma(\hat{\boldsymbol{s}}) [\nabla_s \hat{\boldsymbol{x}}(\boldsymbol{y}, \boldsymbol{s})]' \tag{6}$$

where $\Sigma(\hat{\boldsymbol{s}})$ is the covariance matrix of the estimated scatter sinogram.

To compute $\nabla_s \hat{\boldsymbol{x}}(\boldsymbol{y}, \boldsymbol{s})$, we follow the idea presented in [6]. We restrict our attention to the situations where the solution of (3) satisfies

$$0 = \frac{\partial}{\partial x_j}\left[L(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{s}) - \beta U(\boldsymbol{x})\right]\Big|_{\boldsymbol{x}=\hat{\boldsymbol{x}}(\boldsymbol{y},\boldsymbol{s})}, \quad j = 1, \ldots, M. \tag{7}$$

While this assumption precludes inequality constraints, it should work fine here because of the uniform background. Differentiating (7) with respect to s_i by applying the chain rule and solving

the resulting equation, we get [6]

$$\nabla_s \hat{\boldsymbol{x}}(\boldsymbol{y},\boldsymbol{s}) = \left\{-\nabla_{xx}\left[L(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{s})-\beta U(\boldsymbol{x})\right]|_{\boldsymbol{x}=\hat{\boldsymbol{x}}(\boldsymbol{y},\boldsymbol{s})}\right\}^{-1} \nabla_{xs}\left[L(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{s})-\beta U(\boldsymbol{x})\right]|_{\boldsymbol{x}=\hat{\boldsymbol{x}}(\boldsymbol{y},\boldsymbol{s})} \quad (8)$$

where the $(j,k)^{th}$ element of the operator ∇_{xx} is $\frac{\partial^2}{\partial x_j \partial x_k}$, and the $(j,l)^{th}$ element of the operator ∇_{xs} is $\frac{\partial^2}{\partial x_j \partial s_l}$.

From (1) and (2), we can derive

$$\nabla_{xx}\left[L(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{s})-\beta U(\boldsymbol{x})\right] = \boldsymbol{P}' \operatorname{diag}\left[\frac{y_i}{(\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s}+\boldsymbol{r})_i^2}\right]\boldsymbol{P}+\beta\boldsymbol{R}$$

and

$$\nabla_{xs}\left[L(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{s})-\beta U(\boldsymbol{x})\right] = \boldsymbol{P}' \operatorname{diag}\left[\frac{-y_i}{(\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s})_i^2}\right].$$

Then

$$\nabla_s \hat{\boldsymbol{x}}(\boldsymbol{y},\boldsymbol{s}) = \left\{\boldsymbol{P}' \operatorname{diag}\left[\frac{y_i}{(\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s}+\boldsymbol{r})_i^2}\right]\boldsymbol{P}+\beta\boldsymbol{R}\right\}^{-1} \boldsymbol{P}' \operatorname{diag}\left[\frac{y_i}{(\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s})_i^2}\right]. \quad (9)$$

Substituting (9) into (6) results in

$$\Sigma(\hat{\boldsymbol{x}}) \approx \left\{\boldsymbol{P}' \operatorname{diag}\left[\frac{y_i}{(\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s}+\boldsymbol{r})_i^2}\right]\boldsymbol{P}+\beta\boldsymbol{R}\right\}^{-1} \boldsymbol{P}' \operatorname{diag}\left[\frac{y_i}{(\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s})_i^2}\right]\Sigma_{\hat{\boldsymbol{s}}} \operatorname{diag}\left[\frac{y_i}{(\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s})_i^2}\right]\boldsymbol{P} \left\{\boldsymbol{P}' \operatorname{diag}\left[\frac{y_i}{(\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s}+\boldsymbol{r})_i^2}\right]\boldsymbol{P}+\beta\boldsymbol{R}\right\}^{-1} \quad (10)$$

In general, $\hat{\boldsymbol{x}}$ is a slightly blurred version of $\boldsymbol{x}$, so the projection $\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s}$ is approximately equal to the mean of the data, $\bar{\boldsymbol{y}}$. Therefore, we can simplify the above expression to

$$\Sigma(\hat{\boldsymbol{x}}) \approx [\boldsymbol{F}+\beta\boldsymbol{R}]^{-1}\boldsymbol{P}' \operatorname{diag}\left[\frac{\sigma_{\hat{s}_i}^2}{\bar{y}_i^2}\right]\boldsymbol{P}[\boldsymbol{F}+\beta\boldsymbol{R}]^{-1}, \quad (11)$$

where $\boldsymbol{F} = \boldsymbol{P}' \operatorname{diag}\left[\frac{1}{\bar{y}_i}\right]\boldsymbol{P}$ is the Fisher information matrix and $\sigma_{\hat{s}_i}$ is the variance of $\hat{s}_i$. Eq. (11) is the covariance matrix of the noise in the reconstruction that is propagated from the estimated scatter sinogram.

The covariance of noise in the reconstruction caused by the Poisson noise in the data is [6]

$$\Sigma_{\text{Poisson}}(\hat{\boldsymbol{x}}) \approx \nabla_y \hat{\boldsymbol{x}}(\bar{\boldsymbol{y}})\operatorname{Cov}(\boldsymbol{y})[\nabla_y \hat{\boldsymbol{x}}(\bar{\boldsymbol{y}})]' \approx [\boldsymbol{F}+\beta\boldsymbol{R}]^{-1}\boldsymbol{F}[\boldsymbol{F}+\beta\boldsymbol{R}]^{-1}. \quad (12)$$

Adding (11) and (12), we get the covariance of the total noise

$$\Sigma_{\text{total}}(\hat{\boldsymbol{x}}) = [\boldsymbol{F}+\beta\boldsymbol{R}]^{-1}\boldsymbol{P}' \operatorname{diag}\left[\frac{\sigma_{\hat{s}_i}^2}{\bar{y}_i^2}+\frac{1}{\bar{y}_i}\right]\boldsymbol{P}[\boldsymbol{F}+\beta\boldsymbol{R}]^{-1} \quad (13)$$

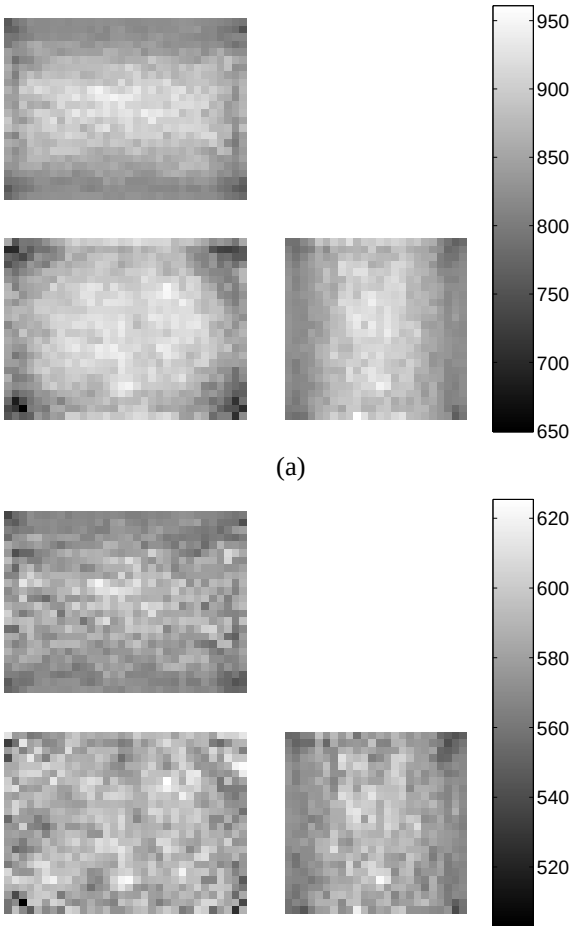

Fig. 2. (a) Reconstruction without scatter correction. (b) Reconstruction with scatter correction. The images are top view slice (upper left), front view slice (lower left), and side view slice (lower right) through the center voxel. Images in (a) and (b) are individually scaled as indicated by the gray level bars.

It shows that the noise from the scatter sinogram is equivalent to an increase of the noise in the data by a factor of $1+\sigma_{\hat{s}_i}^2/\bar{y}_i$. $\sigma_{\hat{s}_i}^2/\bar{y}_i$ is equal to the scatter fraction of the i^{th} LOR divided by the ratio between total number of detection in the Monte Carlo simulation and the total number of detection in the data. For example, if the average scatter fraction is 30%, and the Monte Carlo simulation has 30 times as many events as the data, then the noise variance increase in reconstruction cause by the scatter sinogram will be about 1%. Equation (13) can be used to determine the number of events required in the Monte Carlo simulation and to design better simulation strategy.

III. Simulation Results

In simulation we assume a subject weighing 70 kg and an injection of 1 mCi of FDG, which is uniformly distributed within the body. This activity density within the $72 \times 72 \times 96$ mm^3 field of view and an imaging time of 60 s gives about 16 million disintegrations within the imaging volume.

The Monte Carlo simulation uses the appropriate energy-dependent cross sections for the interaction of photons in water

(in the field of view) and in LSO detector. The average detection efficiency is about 13% for an energy threshold of 270 keV. Of all the detected events, there are about 35% unscattered events, 32% events scattered in the FOV, and 33% events scattered in the detector (not scattered in the FOV).

Fig. 2 shows some example reconstructions of a simulated flood source with and without scatter correction. The images shown are three orthogonal slices through the center voxel: top view slice (upper left), front view slice (lower left), and side view slice (lower right). The reconstructed image without scatter correction (Fig. 2a) shows brighter at the center of FOV and darker at the corners, especially in the front view slice. The scatter corrected image (Fig. 2b) shows more uniform activity distribution. Note the gray level maps in Fig. 2a and Fig. 2b are different. Here we have only corrected for the scatters in the FOV. We are looking for a better way to deconvolve the scatters in the detector as they are more localized.

IV. Conclusion

We have implemented the Monte Carlo scatter correction method for the list mode likelihood reconstruction algorithm for the PEM and shown some results based on computer simulations. For breast imaging with FDG, we can assume that the FOV is filled with uniform activity and that features such as cancerous lesions account for a small fraction of the radioactivity. This specific application of PEM allows us to run the Monte Carlo scatter simulation only once for each scanner configuration. The scatter sinogram for each individual data set can then be estimated using the total number of detections. This saves a large amount of computation time.

We also theoretically analyzed the noise propagation from the estimate scatter sinogram into the final reconstructed image. The results show that the noise propagated from the estimated scatter sinogram is equivalent to increasing the noise variance in each LOR by a factor of $1 + \sigma^2_{\hat{s}_i}/\bar{y}_i$. If we assume the scatter fraction for each LOR is the same, then this factor is a constant for all LORs and it provides a easy way to calculate the required number of events in the Monte Carlo scatter simulation for a given noise level in reconstruction. This noise analysis is applicable to other scatter estimation methods provided that an estimate of the covariance of the estimated scatter sinogram is available.

References

[1] P. Virador, W. Moses, and R. Huesman, "Reconstruction in PET cameras with irregular sampling and depth of interaction capability," *IEEE Transactions on Nuclear Science*, vol. 45, pp. 1225–1230, 1998.

[2] R. Huesman, G. Klein, W. Moses, J. Qi, B. Reutter, and P. Virador, "List mode maximum likelihood reconstruction applied to positron emission mammography with irregular sampling," *IEEE Transactions on Medical Imaging*, vol. 19, pp. 532–537, 2000.

[3] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of Royal Statistical Society, Series B*, vol. 39, pp. 1–38, 1977.

[4] L. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Transactions on Medical Imaging*, vol. 1, pp. 113–122, 1982.

[5] K. Lange and R. Carson, "EM reconstruction algorithms for emission and transmission tomography," *Journal of Computer Assisted Tomography*, vol. 8, pp. 306–316, 1984.

[6] J. Fessler, "Mean and variance of implicitely defined biased estimators (such as penalized maximum likelihood): Applications to tomography," *IEEE Transactions on Image Processing*, vol. 5, pp. 493–506, 1996.